Modern Stochastics: Theory and Applications 2 (2015) 267–286
DOI: 10.15559/15-VMSTA38CNF

Gärtner–Ellis condition for squared asymptotically stationary Gaussian processes

Marina Kleptsyna[a], **Alain Le Breton**[b], **Bernard Ycart**[b,1,*]

[a]*Laboratoire Manceau de Mathématiques, Université du Maine, Avenue Olivier Messiaen, 72085 Le Mans cedex 9, France*
[b]*Laboratoire Jean Kuntzmann, Université Grenoble Alpes, 51 rue des Mathématiques, 38041 Grenoble cedex 9, France*

marina.kleptsyna@univ-lemans.fr (M. Kleptsyna), alain.le-breton@imag.fr (A. Le Breton), bernard.ycart@imag.fr (B. Ycart)

Received: 19 June 2015, Revised: 18 September 2015, Accepted: 18 September 2015, Published online: 2 October 2015

Abstract We establish the Gärtner–Ellis condition for the square of an asymptotically stationary Gaussian process. The same limit holds for the conditional distribution given any fixed initial point, which entails weak multiplicative ergodicity. The limit is shown to be the Laplace transform of a convolution of gamma distributions with Poisson compound of exponentials. A proof based on the Wiener–Hopf factorization induces a probabilistic interpretation of the limit in terms of a regression problem.

Keywords Gärtner–Ellis condition, Gaussian process, Laplace transform
2010 MSC 60G14, 60F10

1 Introduction

The convergence of the scaled cumulant generating functions of a sequence of random variables implies a large deviation principle; this is known as the Gärtner–Ellis condition [6, p. 43]. Our main result is that condition for the square of an asymptotically stationary Gaussian process. Reasons for studying squared Gaussian processes

*Corresponding author.
[1] This research was supported by Laboratoire d'Excellence TOUCAN (Toulouse Cancer).

come from different fields: large deviation theory [19, 5], time series analysis [10], or ancestry-dependent branching processes [16]. Since only nonnegative real-valued random variables are considered here, we shall use logarithms of Laplace transforms instead of cumulant generating functions.

Theorem 1. *Let $(X_t)_{t\in\mathbb{N}}$ be a Gaussian process with mean $m = (m(t))$ and covariance kernel $K = (K(t,s))$: for all $t, s \in \mathbb{Z}$,*

$$\mathbb{E}[X_t] = m(t) \quad \text{and} \quad \mathbb{E}\big[\big(X_t - m(t)\big)\big(X_s - m(s)\big)\big] = K(t,s).$$

Assume:

$$\sup_{t\in\mathbb{Z}} \big|m(t)\big| < +\infty; \tag{H1}$$

$$\sup_{t\geqslant 1} \max_{s=0}^{t-1} \sum_{r=0}^{t-1} \big|K(s,r)\big| < +\infty. \tag{H2}$$

Assume that there exist a constant m_∞ and a positive definite symmetric function k such that:

$$\sum_{t\in\mathbb{Z}} \big|k(t)\big| < \infty; \tag{H3}$$

$$\lim_{t\to+\infty} \frac{1}{t} \sum_{s=0}^{t-1} \big|m(s) - m_\infty\big| = 0; \tag{H4}$$

$$\lim_{t\to+\infty} \frac{1}{t} \sum_{s,r=0}^{t-1} \big|K(s,r) - k(r-s)\big| = 0. \tag{H5}$$

Denote by f the spectral density of k:

$$f(\lambda) = \sum_{t\in\mathbb{Z}} \mathrm{e}^{\mathrm{i}\lambda t} k(t). \tag{1}$$

For $t \geqslant 0$, consider the following Laplace transform:

$$L_t(\alpha) = \mathbb{E}\left[\exp\left(-\alpha \sum_{s=0}^{t-1} X_s^2\right)\right]. \tag{2}$$

Then for all $\alpha \geqslant 0$,

$$\lim_{t\to+\infty} \frac{1}{t} \log\big(L_t(\alpha)\big) = -\ell(\alpha) = -\ell_0(\alpha) - \ell_1(\alpha) \tag{3}$$

with

$$\ell_0(\alpha) = \frac{1}{4\pi} \int_0^{2\pi} \log\big(1 + 2\alpha f(\lambda)\big)\,\mathrm{d}\lambda \tag{4}$$

and

$$\ell_1(\alpha) = m_\infty^2 \alpha \big(1 + 2\alpha f(0)\big)^{-1}. \tag{5}$$

Theorem 1 yields as a particular case the following result of weak multiplicative ergodicity.

Proposition 1. *Under the hypotheses of Theorem 1 and assuming $K(0,0)$ positive, consider*

$$L_{x,t}(\alpha) = \mathbb{E}_x\left[\exp\left(-\alpha \sum_{s=0}^{t-1} X_s^2\right)\right], \tag{6}$$

where $\mathbb{E}_x$ denotes the conditional expectation given $X_0 = x$.
Then for all $\alpha \geqslant 0$ and $x \in \mathbb{R}$,

$$\lim_{t\to+\infty} \frac{1}{t} \log\bigl(L_{x,t}(\alpha)\bigr) = -\ell(\alpha),$$

where ℓ is defined by (3), (4), and (5).

The analogue for finite-state Markov chains has long been known [6, p. 72]. It was extended to strong multiplicative ergodicity of exponentially converging Markov chains by Meyn and his coworkers; see [14]. In [13], the square of a Gauss–Markov process was studied, strong multiplicative ergodicity was proved, and the limit was explicitly computed. This motivated the present generalization.

The particular case of a centered stationary process ($m(t) = 0$, $K(t,s) = k(t-s)$) can be considered as classical: in that case, the limit (4) follows from Szegő's theorem on Toeplitz matrices: see [9, 4] as a general reference on Toeplitz matrices and [2] for a review of probabilistic applications of Szegő's theory. The extension to the centered asymptotically stationary case follows from the notion of asymptotically equivalent matrices in the L^2 sense; see Section 7.4, p. 104, of [9], and [8]. The noncentered stationary case ($m(t) = m_\infty$ and $K(s,t) = k(s-t)$) has received much less attention. In Proposition 2.2 of [5], the large deviation principle is obtained for a squared noncentered stationary Gaussian process. There, the centered case is deduced from Szegő's theorem, whereas the noncentered case follows from the contraction principle. A similar approach to the general case can be found in [1].

We propose here a different method. Instead of the spectral decomposition and Szegő's theorem, a Wiener–Hopf factorization is used. The limits (4) and (5) are both deduced from the asymptotics of that factorization. The technique is close to those developed in [12] and used in [13]. One advantage is that the coefficients of the Wiener–Hopf factorization can be given a probabilistic interpretation in terms of a regression problem. This approach will be detailed in Section 2.

To go from the stationary to the asymptotically stationary case, the asymptotic equivalence of matrices is needed. But the classical L^2 definition of [8, Sect. 2.3] does not suffice for the noncentered case. A stronger notion, linked to the L^1 norm of vectors instead of the L^2 norm, will be developed in Section 3.

Joining the stationary case to asymptotic equivalence, we get the conclusion of Theorem 1, but only for small enough values of α. To deduce that the convergence holds for all $\alpha \geqslant 0$, an extension of Lévy's continuity theorem will be used: if both $(L_t(\alpha))^{1/t}$ and $\mathrm{e}^{-\ell(\alpha)}$ are the Laplace transforms of probability distributions on $\mathbb{R}^+$, then the convergence over an interval implies the weak convergence of measures and hence the convergence of Laplace transforms for all $\alpha \geqslant 0$. In fact, $(L_t(\alpha))^{1/t}$ and

$e^{-\ell(\alpha)}$ both are the Laplace transforms of infinitely divisible distributions, more precisely, convolutions of gamma distributions with Poisson compounds of exponentials. Details will be given in Section 4, together with the particular case of a Gauss–Markov process.

2 The stationary case

This section treats the stationary case: $m(t) = m_\infty$ and $K(s,t) = k(t-s)$. We shall denote by $c_t = (m_\infty)_{s=0,\dots,t-1}$ the constant vector with coordinates all equal to m_∞ and by H_t the Toeplitz matrix with symbol k: $H_t = (k(s-r))_{s,r=0,\dots,t-1}$. The main result of this section is a particular case of Theorem 1. It entails Proposition 2.2 of Bryc and Dembo [5].

Proposition 2. *Assume that k is a positive definite symmetric function such that*

$$\sum_{t\in\mathbb{Z}} \left|k(t)\right| = M < +\infty,$$

and denote by f the corresponding spectral density:

$$f(\lambda) = \sum_{t\in\mathbb{Z}} e^{i\lambda t} k(t).$$

Let $Z = (Z_t)_{t\in\mathbb{Z}}$ be a centered stationary process with covariance function k. Let m_∞ be a real. For all α such that $0 \leqslant \alpha < 1/(2M)$,

$$\lim_{t\to+\infty} \frac{1}{t} \log\left(\mathbb{E}\left[\exp\left(-\alpha \sum_{s=0}^{t-1} (Z_s + m_\infty)^2\right)\right]\right) = -\ell_0(\alpha) - \ell_1(\alpha),$$

where $\ell_0(\alpha)$ and $\ell_1(\alpha)$ are defined by (4) and (5).

Denote by m_t and K_t the mean and covariance matrix of the vector $(X_s)_{s=0,\dots,t-1}$. The Laplace transform of the squared norm of a Gaussian vector has a well-known explicit expression; see, for instance, [19, p. 6]. The identity matrix indexed by $0,\dots,t-1$ is denoted by I_t, and the transpose of a vector m is denoted by m^*. Then

$$L_t(\alpha) = \left(\det(I_t + 2\alpha K_t)\right)^{-1/2} \exp\left(-\alpha m_t^* (I_t + 2\alpha K_t)^{-1} m_t\right), \tag{7}$$

In the stationary case, $m_t = c_t$ and $K_t = H_t$. From (7) we must prove that the following two limits hold:

$$\lim_{t\to+\infty} \frac{1}{2t} \log\left(\det(I_t + 2\alpha H_t)\right) = \ell_0(\alpha) = \frac{1}{4\pi} \int_0^{2\pi} \log\left(1 + 2\alpha f(\lambda)\right) d\lambda \tag{8}$$

and

$$\lim_{t\to+\infty} \frac{\alpha}{t} c_t^* (I_t + 2\alpha H_t)^{-1} c_t = \ell_1(\alpha) = m_\infty^2 \alpha \left(1 + 2\alpha f(0)\right)^{-1}. \tag{9}$$

Here, $I_t + 2\alpha H_t$ will be interpreted as the covariance matrix of the random vector $(Y_s)_{s=0,\dots,t-1}$ from the process

$$Y = \varepsilon + \sqrt{2\alpha} Z, \tag{10}$$

where $\varepsilon = (\varepsilon_t)_{t\in\mathbb{Z}}$ is a sequence of i.i.d. standard normal random variables, independent from Z. The limits (8) and (9) will be deduced from a Cholesky decomposition of $I_t + 2\alpha H_t$. We begin with an arbitrary positive definite matrix A. The Cholesky decomposition writes it as the product of a lower triangular matrix by its transpose. Thus, A^{-1} is the product of an upper triangular matrix by its transpose. Write it as $A^{-1} = T^* D T$, where T is a unit lower triangular matrix (the diagonal coefficients equal to 1), and D is a diagonal matrix with positive coefficients. Denote by G the lower triangular matrix DT. Then $GA = (T^*)^{-1}$ is a unit upper triangular matrix. Hence, the coefficients $G(s,r)$ of G are uniquely determined by the following system of linear equations: for $0 \leqslant s \leqslant t$,

$$\sum_{r=0}^{t} G(t,r)\, A(r,s) = \delta_{t,s}, \tag{11}$$

where $\delta_{t,s}$ denotes the Kronecker symbol equal to 1 if $t = s$ and 0 else. Notice that $A^{-1} = G^* D^{-1} G$, and $TAT^* = D^{-1}$, where D is the diagonal matrix with diagonal entries $G(s,s)$. In particular,

$$\det(\mathrm{A}) = \left(\prod_s G(s,s)\right)^{-1}, \tag{12}$$

and for any vector $m = (m(r))$,

$$m^* A^{-1} m = \sum_s \frac{1}{G(s,s)} \left(\sum_{r=0}^{s} G(s,r) m(r)\right)^2. \tag{13}$$

Here is the probabilistic interpretation of the coefficients $G(t,s)$. Consider a centered Gaussian vector Y with covariance matrix A. For $t = 0,\dots,n$, denote by $\mathcal{Y}_{[\![0,t]\!]}$ the σ-algebra generated by $Y_0,\dots,Y_t$, and by ν_t the partial innovation

$$\nu_t = Y_t - \mathbb{E}[Y_t \mid \mathcal{Y}_{[\![0,t-1]\!]}]$$

with the convention $\nu_0 = Y_0$. Using elementary properties of Gaussian vectors, it is easy to check that

$$\nu_t = \frac{1}{G(t,t)} \sum_{r=0}^{t} G(t,r)\, Y_r. \tag{14}$$

Moreover, the ν_t are independent, and the variance of ν_t is $1/G(t,t)$.

When this is applied to $A = I_t + 2\alpha H_t$, another interesting interpretation arises. For $t = 0,\dots,n$, $(G(t,s))_{s=0,\dots,t}$ is the unique solution to the system

$$G(t,s) + 2\alpha \sum_{r=0}^{t} G(t,r)\, k(r-s) = \delta_{t,s}. \tag{15}$$

Observe that Eqs. (15) are the normal equations of the regression of the ε_t over the Y_t in the model (10). Actually, since $\mathbb{E}[Y_r Y_s] = \delta_{s,r} + 2\alpha k(r-s)$ and $\mathbb{E}[\varepsilon_t Y_s] = \delta_{t,s}$, setting

$$\mu_t = G(t,t)\,\nu_t = \sum_{r=0}^{t} G(t,r)\,Y_r, \tag{16}$$

Eq. (15) says that for $s = 0, \ldots, t$,

$$\mathbb{E}[\mu_t Y_s] = \mathbb{E}[\varepsilon_t Y_s].$$

This means that

$$\mu_t = \mathbb{E}[\varepsilon_t \mid \mathcal{Y}_{[\![0,t]\!]}\,].$$

Obviously, the μ_t are independent, the variance of μ_t is $G(t,t)$, and the filtering error is

$$\mathbb{E}\big[(\varepsilon_t - \mu_t)^2\big] = 1 - G(t,t).$$

In particular, it follows that $0 < G(t,t) < 1$.

The asymptotics of $G(t,s)$ will now be related to the spectral density f. Denote $g_t(s) = G(t, t-s)$. A change of index in (15) shows that $(g_t(s))_{s=0,\ldots,t}$ is the unique solution to the system

$$g_t(s) + 2\alpha \sum_{r=0}^{t} g_t(r)\,k(s-r) = \delta_{s,0}. \tag{17}$$

Proposition 3. *Assume that k is a positive definite symmetric function such that*

$$\sum_{t\in\mathbb{Z}} \big|k(t)\big| = M < +\infty,$$

and denote by f the corresponding spectral density:

$$f(\lambda) = \sum_{t\in\mathbb{Z}} \mathrm{e}^{\mathrm{i}\lambda t} k(t).$$

For all α such that $0 \leqslant \alpha < 1/(2M)$, the following equation has a unique solution in $L^1(\mathbb{Z})$:

$$g(s) + 2\alpha \sum_{r=0}^{+\infty} g(r)\,k(s-r) = \delta_{s,0}. \tag{18}$$

We have:

$$g(0) = \exp\left(-\frac{1}{2\pi}\int_0^{2\pi} \log\big(1 + 2\alpha f(\lambda)\big)\,\mathrm{d}\lambda\right) \tag{19}$$

and

$$\sum_{s=0}^{+\infty} g(s) = \exp\left(-\frac{1}{2}\log\big(1 + 2\alpha f(0)\big) - \frac{1}{4\pi}\int_0^{2\pi} \log\big(1 + 2\alpha f(\lambda)\big)\,\mathrm{d}\lambda\right). \tag{20}$$

Moreover, if $g_t(s)$ *is defined for all* $0 \leqslant s \leqslant t$ *by* (17)*, then for all* $s \geqslant 0$,

$$\lim_{t\to+\infty} g_t(s) = g(s) \tag{21}$$

and

$$\lim_{t\to+\infty} \sum_{s=0}^{t} g_t(s) = \sum_{s=0}^{+\infty} g(s). \tag{22}$$

The proof is equivalent to writing the Wiener–Hopf factorization of the operator $I + 2\alpha H$: compare with Section 1.5 of [4], in particular, with the proof of Theorem 1.14 on p. 17. The main idea is to reduce Eq. (18) to the problem of finding a sectionally holomorphic function satisfying a boundary condition on the unit circle. This idea is originally due to Krein [15].

Proof. Conditions of invertibility for Toeplitz operators are well known. They are treated in Sections 2.3 and 7.2 of [4]. Here, the L^1 norm of the Toeplitz operator H with symbol k is M, and the condition $0 \leqslant \alpha < 1/(2M)$ permits to write the inverse as

$$(I + 2\alpha H)^{-1} = \sum_{n=0}^{+\infty} (-2\alpha H)^n.$$

This property implies the existence and uniqueness of the solution to Eq. (18). The convergence of the truncated inverse $(I_t + 2\alpha H_t)^{-1}$ to $(I + 2\alpha H)^{-1}$ is deduced for the L^2 case from [4, p. 42]. The convergence of entries follows, and hence (21). To obtain (22), consider $\Delta_t(s) = g(s) - g_t(s)$. From (17) and (18) we have

$$\Delta_t(s) = -2\alpha \sum_{r=0}^{t} k(r-s)\Delta_t(r) - 2\alpha \sum_{r=t+1}^{+\infty} g(r)k(r-s).$$

Hence,

$$\begin{aligned}\sum_{s=0}^{t} \left|\Delta_t(s)\right| &\leqslant 2\alpha \Bigg(\sum_{r=0}^{t} \left|\Delta_t(r)\right|\Bigg)\Bigg(\sum_{s=-\infty}^{+\infty} \left|k(s)\right|\Bigg) \\ &\quad + 2\alpha \Bigg(\sum_{s=-\infty}^{+\infty} \left|k(s)\right|\Bigg)\Bigg(\sum_{r=t+1}^{+\infty} \left|g(r)\right|\Bigg).\end{aligned}$$

Thus, we obtain the following bound:

$$\sum_{s=0}^{t} \left|\Delta_t(s)\right| \leqslant \frac{2\alpha M}{1 - 2\alpha M} \sum_{r=t+1}^{+\infty} \left|g(r)\right|,$$

which yields (22).

Now we prove identities (19) and (20). The generating function of $(g(s))_{s\geqslant 0}$ will be first related to the spectral density f. Define for all $s \in \mathbb{Z}$,

$$g^+(s) = \begin{cases} g(s) & \text{if } s \geqslant 0, \\ 0 & \text{else,} \end{cases} \quad \text{and} \quad g^-(s) = \begin{cases} g(s) & \text{if } s < 0, \\ 0 & \text{else.} \end{cases}$$

Denote by F^+ and F^- the Fourier transforms of g^+ and g^-:

$$F^{\pm}(\lambda) = \sum_{s\in\mathbb{Z}} \mathrm{e}^{\mathrm{i}s\lambda} g^{\pm}(s).$$

Take the Fourier transforms in both members of (18):

$$F^+(\lambda) + F^-(\lambda) + 2\alpha F^+(\lambda) f(\lambda) = 1$$

or

$$F^+(\lambda)\big(1 + 2\alpha f(\lambda)\big) = 1 - F^-(\lambda). \tag{23}$$

Let us define the sectionally holomorphic function φ as follows (see [18]):

$$\varphi(\zeta) = \begin{cases} \varphi^+(\zeta) = \sum_{s\geqslant 0} \zeta^s g^+(s) & \text{if } |\zeta| < 1, \\ \varphi^-(\zeta) = 1 - \sum_{s<0} \zeta^s g^-(s) & \text{if } |\zeta| > 1. \end{cases}$$

Then:

$$F^+(\lambda) = \lim_{\substack{\zeta\to \mathrm{e}^{\mathrm{i}\lambda} \\ |\zeta|<1}} \varphi(\zeta)$$
$$F^-(\lambda) = \lim_{\substack{\zeta\to \mathrm{e}^{\mathrm{i}\lambda} \\ |\zeta|>1}} 1 - \varphi(\zeta),$$

and Eq. (23) expresses the boundary condition

$$\varphi^+(\zeta) = \frac{1}{1 + 2\alpha \widetilde{f}(\zeta)} \varphi^-(\zeta), \quad |\zeta| = 1, \tag{24}$$

where $\widetilde{f}(\zeta)$ denotes the value of $f(\lambda)$ for $\zeta = \mathrm{e}^{\mathrm{i}\lambda}$. Problem (24) is a well-known homogeneous Riemann problem. Since by construction φ is bounded near infinity and for $|\zeta| = 1$, $1 + 2\alpha\widetilde{f}(\zeta) > 0$, the solution of (24) can be written explicitly [18, §35]. Assuming for a moment that $\widetilde{f}$ satisfies the Hölder condition on the unit circle, we have that for all ζ_0,

$$\varphi(\zeta_0) = \exp\left(-\frac{1}{2\pi\mathrm{i}} \oint_{|\zeta|=1} \frac{\log(1 + 2\alpha\widetilde{f}(\zeta))}{\zeta - \zeta_0}\, \mathrm{d}\zeta\right). \tag{25}$$

Observe that the choice of a branch for the logarithm does not change the result. From now on, the principal branch will be taken.

Equation (25) for $\zeta_0 = 0$ implies immediately that

$$\begin{aligned} g^+(0) &= \exp\left(-\frac{1}{2\pi\mathrm{i}} \oint_{|\zeta|=1} \frac{\log(1 + 2\alpha\widetilde{f}(\zeta))}{\zeta}\, \mathrm{d}\zeta\right) \\ &= \exp\left(-\frac{1}{2\pi} \int_0^{2\pi} \log\big(1 + 2\alpha f(\lambda)\big)\, \mathrm{d}\lambda\right), \end{aligned}$$

which is (19).

To prove (20), we will calculate

$$\begin{aligned}
&\lim_{\substack{\zeta_0\to 1\\ |\zeta_0|<1}} -\frac{1}{2\pi\mathrm{i}}\oint_{|\zeta|=1}\frac{\log(1+2\alpha\widetilde{f}(\zeta))}{\zeta-\zeta_0}\,\mathrm{d}\zeta\\
&\quad=\lim_{\substack{\zeta_0\to 1\\ |\zeta_0|<1}} -\frac{1}{2\pi\mathrm{i}}\oint_{|\zeta|=1}\frac{\log(1+2\alpha\widetilde{f}(1))}{\zeta-\zeta_0}\,\mathrm{d}\zeta\\
&\qquad-\lim_{\substack{\zeta_0\to 1\\ |\zeta_0|<1}} \frac{1}{2\pi\mathrm{i}}\oint_{|\zeta|=1}\frac{\log(1+2\alpha\widetilde{f}(\zeta))-\log(1+2\alpha\widetilde{f}(1))}{\zeta-\zeta_0}\,\mathrm{d}\zeta.
\end{aligned}$$

The first integral does not depend on ζ_0: it is equal to

$$-\log\bigl(1+2\alpha\widetilde{f}(1)\bigr)=-\log\bigl(1+2\alpha f(0)\bigr). \tag{26}$$

Still assuming that $\widetilde{f}$ satisfies a Hölder condition on the unit circle, the second limit exists and is equal to Cauchy's principal value integral [18]:

$$\begin{aligned}
&\frac{1}{2\pi\mathrm{i}}\oint_{|\zeta|=1}\frac{\log(1+2\alpha\widetilde{f}(\zeta))-\log(1+2\alpha\widetilde{f}(1))}{\zeta-1}\,\mathrm{d}\zeta\\
&\quad=\lim_{\epsilon\to 0}\frac{1}{2\pi\mathrm{i}}\oint_{\substack{|\zeta|=1\\ |\arg(\zeta)|>\epsilon}}\frac{\log(1+2\alpha\widetilde{f}(\zeta))-\log(1+2\alpha\widetilde{f}(1))}{\zeta-1}\,\mathrm{d}\zeta\\
&\quad=\lim_{\epsilon\to 0}\frac{1}{2\pi}\int_{\substack{[-\pi,\pi]\\ |\lambda|>\epsilon}}\frac{\log(1+2\alpha f(\lambda))-\log(1+2\alpha f(0))}{1-\mathrm{e}^{\mathrm{i}\lambda}}\,\mathrm{d}\lambda.
\end{aligned}$$

Now for $\epsilon<|\lambda|<\pi$,

$$\frac{1}{1-\mathrm{e}^{-\mathrm{i}\lambda}}=\frac{1}{2}+\mathrm{i}\frac{\sin(\lambda)}{2(1-\cos(\lambda))}.$$

The imaginary part is an odd function of λ, which is multiplied by an even function inside the integral. Hence, the imaginary part in the last integral vanishes. Therefore,

$$\begin{aligned}
&\frac{1}{2\pi\mathrm{i}}\oint_{|\zeta|=1}\frac{\log(1+2\alpha\widetilde{f}(\zeta))-\log(1+2\alpha\widetilde{f}(1))}{\zeta-1}\,\mathrm{d}\zeta\\
&\quad=-\frac{1}{2}\log\bigl(1+2\alpha f(0)\bigr)+\frac{1}{4\pi}\int_{-\pi}^{\pi}\log\bigl(1+2\alpha f(\lambda)\bigr)\,\mathrm{d}\lambda.
\end{aligned}$$

Substracting the last equation from (26) and taking exponential, we get

$$\begin{aligned}
\varphi^+(1)&=\lim_{\substack{\zeta\to 1\\ |\zeta|<1}}\varphi(\zeta)\\
&=\exp\biggl(-\frac{1}{2}\log\bigl(1+2\alpha f(0)\bigr)-\frac{1}{4\pi}\int_{0}^{2\pi}\log\bigl(1+2\alpha f(\lambda)\bigr)\,\mathrm{d}\lambda\biggr),
\end{aligned}$$

which is (20).

To finish the proof, we must explain how the extra Hölder condition on $\widetilde{f}$ can be removed. It must be emphasized here that the problem is not to obtain the solution of a Riemann problem without Hölder condition on the boundary, but only the values of $\varphi(0)$ and $\varphi^+(1)$. For this, a truncation argument can be used. From the covariance function k on $\mathbb{Z}$, define

$$k_N(s) = \begin{cases} k(s) & \text{if } |s| \leqslant N, \\ 0 & \text{else.} \end{cases}$$

Replace k by k_N in (18) and denote the solution by g_N. The spectral density f_N, which is the Fourier transform of k_N, is smooth. Therefore, the Hölder condition on the unit circle is satisfied for $\widetilde{f}_N$. The previous proof shows that Eqs. (19) and (20) hold for g_N and f_N. But g_N converges to g in $L^1(\mathbb{Z})$, and f_N converges uniformly to f. Taking the limit in N yields the desired result. □

Here is the probabilistic interpretation. Consider a centered stationary process $(Y_t)_{t\in\mathbb{Z}}$ with covariance function $A(t,s) = a(t-s)$. For $s \leqslant t$, denote by $\mathcal{Y}_{[\![s,t]\!]}$ the σ-algebra generated by $(Y_r)_{r=s,\dots,t}$. Consider again the partial innovation $\nu_t = Y_t - \mathbb{E}[Y_t \mid \mathcal{Y}_{[\![0,t-1]\!]}]$. From (14) and using stationarity, ν_t has the same distribution as

$$\eta_t = \frac{1}{G(t,t)} \sum_{r=0}^{t} G(t, t-r)\, Y_{-r},$$

which is

$$\eta_t = Y_0 - \mathbb{E}[Y_0 \mid \mathcal{Y}_{[\![-t,-1]\!]}].$$

As t tends to infinity, η_t converges almost surely to

$$\eta_\infty = Y_0 - \mathbb{E}[Y_0 \mid \mathcal{Y}_{[\![-\infty,-1]\!]}].$$

Observe by stationarity that for all r,

$$\eta_\infty \stackrel{\mathcal{D}}{=} Y_r - \mathbb{E}[Y_r \mid \mathcal{Y}_{[\![-\infty,r-1]\!]}],$$

which is the innovation process associated to Y. Now the variance of ν_t, $1/G(t,t)$, tends to the variance of η_∞. By the Szegő–Kolmogorov formula (see, e.g., Theorem 3 on p. 137 of [10]) that variance is

$$\exp\left(\frac{1}{2\pi} \int_0^{2\pi} \log\bigl(\phi(\lambda)\bigr)\, \mathrm{d}\lambda\right),$$

where $\phi(\lambda)$ is the spectral density of Y. Let X be a centered stationary process with covariance function k, ε be a standard Gaussian noise, and $Y = \varepsilon + \sqrt{2\alpha}X$. The spectral densities ϕ of Y and f of X are related by $\phi(\lambda) = 1 + 2\alpha f(\lambda)$. Hence,

$$\lim_{t\to+\infty} \operatorname{var}(\nu_t) = \lim_{t\to+\infty} \frac{1}{G(t,t)} = \exp\left(\frac{1}{2\pi} \int_0^{2\pi} \log\bigl(1 + 2\alpha f(\lambda)\bigr)\, \mathrm{d}\lambda\right),$$

which is equivalent to (19).

Alternatively, observe that, due to stationarity, μ_t defined by (16) has the same distribution as

$$\xi_t = \sum_{r=0}^{t} G(t, t-r)\, Y_{-r},$$

which is

$$\xi_t = \mathbb{E}[\varepsilon_0 \mid \mathcal{Y}_{[\![-t,0]\!]}].$$

As t tends to infinity, ξ_t converges a.s. to

$$\xi_\infty = \mathbb{E}[\varepsilon_0 \mid \mathcal{Y}_{[\![-\infty,0]\!]}].$$

Of course, since $\mathbb{E}[\varepsilon_{-s} Y_{-r}] = \delta_{s,r}$ for all $s = 0, \dots, t$,

$$\mathbb{E}[\xi_t\, \varepsilon_{-s}] = G(t, t-s).$$

Hence, the limiting property (21) says that

$$\mathbb{E}[\xi_\infty\, \varepsilon_{-s}] = \lim_{t\to+\infty} G(t, t-s) = g(s).$$

In fact, ξ_∞ admits the representation

$$\xi_\infty = \sum_{s=0}^{+\infty} g(s)\, Y_{-s}.$$

Similarly, for all t,

$$\mathbb{E}[\varepsilon_t \mid \mathcal{Y}_{[\![-\infty,t]\!]}] = \sum_{s=0}^{+\infty} g(s)\, Y_{t-s},$$

which means that $(g(s))$ realizes the optimal causal Wiener filter of ε_t from the Y_{t-s}.

Now, Proposition 2 is a straightforward consequence of Proposition 3.

Proof. Let the coefficients $g_\tau(s)$ be defined by (17). Applying (12) to $A = I_t + 2\alpha H_t$, we get

$$\big(\det(I_t + 2\alpha H_t)\big)^{-1/2} = \left(\prod_{\tau=0}^{t-1} g_\tau(0)\right)^{1/2}.$$

Therefore,

$$\frac{1}{t}\log\Big(\big(\det(I_t + 2\alpha H_t)\big)^{-1/2}\Big) = \frac{1}{2t}\sum_{\tau=0}^{t-1}\log\big(g_\tau(0)\big).$$

From Proposition 3 we have

$$\lim_{\tau\to+\infty} g_\tau(0) = g(0) = \exp\left(-\frac{1}{2\pi}\int_0^{2\pi}\log\big(1 + 2\alpha f(\lambda)\big)\,\mathrm{d}\lambda\right).$$

Hence,

$$\lim_{t\to+\infty}\frac{1}{t}\log\Big(\big(\det(I_t + 2\alpha H_t)\big)^{-1/2}\Big) = -\ell_0(\alpha).$$

Applying now (13) to $A = I_t + 2\alpha H_t$, we get

$$c_t^* G_t^* D_t^{-1} G_t c_t = m_\infty^2 \sum_{\tau=0}^{t-1} \frac{1}{g_\tau(0)} \left(\sum_{s=0}^{\tau} g_\tau(s) \right)^2 .$$

From Proposition 3 we have

$$\lim_{\tau \to +\infty} \frac{1}{g_\tau(0)} \left(\sum_{s=0}^{\tau} g_\tau(s) \right)^2 = \frac{1}{g(0)} \left(\sum_{s=0}^{+\infty} g(s) \right)^2$$
$$= \bigl(1 + 2\alpha f(0)\bigr)^{-1}.$$

Hence,

$$\lim_{t \to +\infty} \frac{\alpha}{t} c_t^* (I_t + 2\alpha H_t)^{-1} c_t = \ell_1(\alpha). \qquad \square$$

3 Asymptotic equivalence

Proposition 2 only treats the stationary case. To extend the result under the hypotheses of Theorem 1, a notion of asymptotic equivalence of matrices and vectors is needed. It is developed in this section.

From (7), we must prove that, under the hypotheses of Theorem 1,

$$\lim_{t \to +\infty} \frac{1}{2t} \log\bigl(\det(I_t + 2\alpha K_t)\bigr) = \ell_0(\alpha) = \frac{1}{4\pi} \int_0^{2\pi} \log\bigl(1 + 2\alpha f(\lambda)\bigr)\, \mathrm{d}\lambda \tag{27}$$

and

$$\lim_{t \to +\infty} \frac{\alpha}{t} m_t^* (I_t + 2\alpha K_t)^{-1} m_t = \ell_1(\alpha) = m_\infty^2 \alpha \bigl(1 + 2\alpha f(0)\bigr)^{-1}. \tag{28}$$

If $K_t = H_t$ (centered stationary case), then (27) is (8). It can also be obtained by a straightforward application of Szegő's theorem; see [4, 2]. Relation (27) (centered asymptotically stationary case) is a consequence of the theory of asymptotically Toeplitz matrices; see Section 7.4 on p. 104 of [9] and also [8, Theorem 4 on p. 178]. Asymptotic equivalence of matrices in Szegő's theory is taken in the L^2 sense, which is weaker than that considered here. In other words, (27) holds under weaker hypotheses than (H1–H5). In order to prove (28), we shall develop asymptotic equivalence of matrices and vectors along the same lines as [8, Sect. 2.3], but in a stronger sense, replacing L^2 by L^∞ and L^1, for boundedness and convergence. The norms used here for a vector $v = (v(s))_{s=0,\dots,t-1}$ are

$$\|v\|_\infty = \max_{s=0}^{t-1} \bigl|v(s)\bigr| \quad \text{and} \quad \|v\|_1 = \sum_{s=0}^{t-1} \bigl|v(s)\bigr|.$$

For symmetric matrices, the norm subordinate to $\|\cdot\|_\infty$ is equal to the norm subordinate to $\|\cdot\|_1$. It will be denoted by $\|\cdot\|$ and referred to as the *strong norm*. For $A = (A(s,r))_{s,r=0,\dots,t-1}$ such that $A^* = A$,

$$\|A\| = \max_{s=0}^{t-1} \sum_{r=0}^{t-1} \bigl|A(s,r)\bigr| = \max_{\|v\|_\infty = 1} \|Av\|_\infty$$

$$= \max_{r=0}^{t-1} \sum_{s=0}^{t-1} \left|A(s,r)\right| = \max_{\|v\|_1=1} \|Av\|_1.$$

The following *weak norm* will be denoted by $|A|$:

$$|A| = \frac{1}{t} \sum_{s,r=0}^{t-1} \left|A(s,r)\right|.$$

Clearly, $|A| \leqslant \|A\|$. Moreover, the following bounds hold.

Lemma 1. *Let A and B be two symmetric matrices. Then*

$$|AB| \leqslant \|A\|\,|B| \quad \textit{and} \quad |AB| \leqslant |A|\,\|B\|.$$

Proof. $|AB|$ is the arithmetic mean of the L^1 norms of column vectors of AB. If b is any column vector of B, then

$$\|Ab\|_1 \leqslant \|A\|\,\|b\|_1$$

because the strong norm is subordinate to the L^1 norm of vectors. Hence, the first result. For the second result, replace columns by rows. □

Here is a definition of asymptotic equivalence for vectors.

Definition 1. Let $(v_t)_{t\geqslant 0}$ and $(w_t)_{t\geqslant 0}$ be two sequences of vectors such that for all $t \geqslant 0$, $v_t = (v_t(s))_{s=0,\dots,t-1}$ and $w_t = (w_t(s))_{s=0,\dots,t-1}$. They are said to be *asymptotically equivalent* if:

1. $\|v_t\|_\infty$ and $\|w_t\|_\infty$ are uniformly bounded,
2. $\lim_{t\to+\infty} \frac{1}{t}\|v_t - w_t\|_1 = 0$.

The asymptotic equivalence of (v_t) and (w_t) will be denoted by $v_t \sim w_t$.

Hypotheses (H1) and (H4) imply that $m_t \sim c_t$.

Asymptotic equivalence for matrices is defined as follows (compare with [8, p. 172]).

Definition 2. Let $(A_t)_{t\geqslant 0}$ and $(B_t)_{t\geqslant 0}$ be two sequences of symmetric matrices, where for all $t \geqslant 0$, $A_t = (A_t(s,r))_{s,t=0,\dots,t-1}$ and $B_t = (B_t(s,r))_{s,t=0,\dots,t-1}$. They are said to be *asymptotically equivalent* if:

1. $\|A_t\|$ and $\|B_t\|$ are uniformly bounded,
2. $\lim_{t\to+\infty} |A_t - B_t| = 0$.

The asymptotic equivalence of (A_t) and (B_t) will still be denoted by $A_t \sim B_t$.

Here are some elementary results, analogous to those stated in Theorem 1 on p. 172 of [8].

Lemma 2. *Let (A_t), (B_t), (C_t), (D_t) be four sequences of symmetric matrices.*

1. *If $A_t \sim B_t$ and $B_t \sim C_t$, then $A_t \sim C_t$.*
2. *If $A_t \sim B_t$ and $C_t \sim D_t$, then $A_t + C_t \sim B_t + D_t$.*
3. *If $A_t \sim B_t$ and $C_t \sim D_t$, then $A_t C_t \sim B_t D_t$.*
4. *If $A_t \sim B_t$ and F is an analytic function with radius R such that $R > \max \|A_t\|, \max \|B_t\|$, then $F(A_t) \sim F(B_t)$.*

Proof. Points 1 and 2 follow from the triangle inequality for the weak norm. For point 3, because $\| \cdot \|$ is a norm of matrices, $\|A_t C_t\| \leqslant \|A_t\| \, \|C_t\|$, and $\|B_t D_t\| \leqslant \|B_t\| \, \|D_t\|$ are uniformly bounded. Moreover by Lemma 1,

$$\begin{aligned} |A_t C_t - B_t D_t| &\leqslant \left|(A_t - B_t) C_t\right| + \left|B_t (C_t - D_t)\right| \\ &\leqslant |A_t - B_t| \, \|C_t\| + \|B_t\| \, |C_t - D_t|. \end{aligned}$$

Since $\|C_t\|$ and $\|B_t\|$ are uniformly bounded and

$$\lim_{t \to \infty} |A_t - B_t| = \lim_{t \to \infty} |C_t - D_t| = 0,$$

the result follows. For point 4, let F be analytic with radius of convergence R. For $|z| < R$, let

$$F(z) = \sum_{k=0}^{+\infty} a_k \, z^k$$

and

$$F_n(z) = \sum_{k=0}^{n} a_k \, z^k.$$

The matrices $F(A_t)$, $F(B_t)$ are defined as the limits of $F_n(A_t)$, $F_n(B_t)$; from the hypothesis it follows that the convergence is uniform in t. Because $\| \cdot \|$ is a matrix norm, $\|F(A_t)\| \leqslant F(\|A_t\|)$, and the same holds for B_t: $\|F(A_t)\|$ and $\|F(B_t)\|$ are uniformly bounded. Let ϵ be a positive real. Fix n such that for all t,

$$\left\|F(A_t) - F_n(A_t)\right\| < \frac{\epsilon}{3} \quad \text{and} \quad \left\|F(B_t) - F_n(B_t)\right\| < \frac{\epsilon}{3}.$$

By induction on n using points 2 and 3, $F_n(A_t) \sim F_n(B_t)$. There exists t_0 such that for all $t > t_0$,

$$\left|F_n(A_t) - F_n(B_t)\right| < \frac{\epsilon}{3}.$$

Thus, for all $t > t_0$,

$$\begin{aligned} \left|F(A_t) - F(B_t)\right| &\leqslant \left|F(A_t) - F_n(A_t)\right| + \left|F_n(A_t) - F_n(B_t)\right| + \left|F_n(B_t) - F(B_t)\right| \\ &\leqslant \left\|F(A_t) - F_n(A_t)\right\| + \left|F_n(A_t) - F_n(B_t)\right| + \left\|F_n(B_t) - F(B_t)\right\| \\ &< \epsilon. \end{aligned}$$

Hence the result. $\square$

Hypothesis (H3) implies that $\|H_t\|$ is uniformly bounded, (H2) and (H5) that $K_t \sim H_t$. Point 4 will be applied to $F(z) = (1 + 2\alpha z)^{-1}$, which has the radius of convergence $R = 1/2\alpha$. Let M be defined as

$$M = \max\left\{\max_{t\geqslant 1} \|K_t\|, \sum_{t\in\mathbb{Z}} \left|k(t)\right|\right\}.$$

For all $\alpha < \alpha_0 = 1/(2M)$,

$$(I_t + 2\alpha K_t)^{-1} \sim (I_t + 2\alpha H_t)^{-1}. \tag{29}$$

Here is the relation between asymptotic equivalence of vectors and matrices.

Lemma 3. *1. If $A_t \sim B_t$ and $\|v_t\|_\infty$ is uniformly bounded, then $A_t v_t \sim B_t v_t$.*

2. If $v_t \sim w_t$ and $\|A_t\|$ is uniformly bounded, then $A_t v_t \sim A_t w_t$.

Proof. The norms $\|A_t v_t\|_\infty$, $\|B_t v_t\|_\infty$, $\|A_t w_t\|_\infty$ are uniformly bounded because of the fact that $\|\cdot\|$ is subordinate to $\|\cdot\|_\infty$. Next, for point 1,

$$\frac{1}{t}\left\|(A_t - B_t)v_t\right\|_1 \leqslant \|v_t\|_\infty |A_t - B_t|.$$

For point 2,

$$\frac{1}{t}\left\|A_t(v_t - w_t)\right\|_1 \leqslant \frac{1}{t}\|A_t\|\|v_t - w_t\|_1. \qquad \square$$

The relation between asymptotic equivalence of vectors and our goal is the following.

Lemma 4. *If $v_t \sim w_t$ and $u_t \sim z_t$, then*

$$\lim_{t\to+\infty} \frac{1}{t}\left(v_t^* u_t - w_t^* z_t\right) = 0.$$

Proof.

$$\begin{aligned}\frac{1}{t}\left|v_t^* u_t - w_t^* z_t\right| &\leqslant \frac{1}{t}\left(\left|v_t^*(u_t - z_t)\right| + \left|\left(v_t^* - w_t^*\right)z_t\right|\right)\\ &\leqslant \frac{1}{t}\left(\|v_t\|_\infty \|u_t - z_t\|_1 + \|z_t\|_\infty \|v_t - w_t\|_1\right).\end{aligned}$$

Hence the result. $\square$

Using asymptotic equivalence, (27) and (28) can easily be deduced from (8) and (9) for $0 < \alpha < 1/(2M)$. We shall not detail the passage from (8) to (27); see Theorem 4 on p. 178 of [8]. Here is the passage from (9) to (28). For all $\alpha < 1/(2M)$, it follows from (29) by point 1 of Lemma 3 that

$$(I_t + 2\alpha K_t)^{-1} c_t \sim (I_t + 2\alpha H_t)^{-1} c_t.$$

By point 2 of Lemma 3, we have

$$(I_t + 2\alpha K_t)^{-1} m_t \sim (I_t + 2\alpha H_t)^{-1} c_t.$$

Lemma 4 implies

$$\lim_{t \to +\infty} \frac{1}{t} m_t^* (I_t + 2\alpha K_t) m_t = \lim_{t \to +\infty} \frac{1}{t} c_t^* (1 + 2\alpha H_t)^{-1} c_t.$$

Hence (28).

Still using asymptotic equivalence, it will now be shown that Proposition 1 is just a particular case of Theorem 1. Indeed, consider the Gaussian process X^x with mean

$$m_x(t) = \mathbb{E}\big[X_t^x\big] = m(t) + \frac{K(0,t)}{K(0,0)}\big(x - m(0)\big) \tag{30}$$

and covariance function

$$K^\bullet(t,s) = \mathbb{E}\big[\big(X_t^x - m_x(t)\big)\big(X_s^x - m_x(s)\big)\big] = K(t,s) - \frac{K(t,0)K(s,0)}{K(0,0)}. \tag{31}$$

The distribution of $(X_t^x)_{t\in\mathbb{N}}$ and the conditional distribution of $(X_t)_{t\in\mathbb{N}}$ given $X_0 = x$ are the same. Denote by $m_{x,t}$ and $K_t^\bullet$ the mean and covariance matrix of $(X_s^x)_{s=0,\dots,t-1}$. Theorem 1 applies to X^x, provided that it is proved that $m_{x,t} \sim c_t$ and $K_t^\bullet \sim H_t$. By (H1) and (H2), $\|m_{x,t}\|_\infty$ is uniformly bounded. Moreover, by (30),

$$\frac{1}{t}\|m_{x,t} - m_t\|_1 \leqslant \frac{|x| + \|m_t\|_\infty}{tK(0,0)} \sum_{s=0}^{t-1} \big|K(0,s)\big| \leqslant \frac{|x| + \|m_t\|_\infty}{tK(0,0)} \|K_t\|;$$

thus, $m_{x,t} \sim m_t$, and hence $m_{x,t} \sim c_t$ by transitivity. Now from (31) we have

$$\|K_t^\bullet\| \leqslant \|K_t\| + \max_{r=0}^{t-1} \frac{|K(0,r)|}{K(0,0)} \sum_{s=0}^{t-1} K(0,s) \leqslant \|K_t\| + \frac{\|K_t\|^2}{K(0,0)}.$$

Moreover,

$$\big|K_t^\bullet - K_t\big| \leqslant \frac{1}{tK(0,0)} \Bigg(\sum_{s=0}^{t-1} \big|K(0,s)\big|\Bigg)^2 \leqslant \frac{\|K_t\|^2}{tK(0,0)};$$

thus, $K_t^\bullet \sim K_t$, and hence $K_t^\bullet \sim H_t$ by transitivity (point 1 of Lemma 2).

4 Asymptotic distributions

The results of the two previous sections establish that the conclusion of Theorem 1 holds for a small enough α. To finish the proof, the convergence must be extended to all $\alpha \geqslant 0$. The following variant of Lévy's continuity theorem applies (see Chapter 4 of [11] and, in particular, Exercise 9 on p. 78).

Lemma 5. *Let $\pi, \pi_1, \pi_2, \dots,$ be probability measures on $\mathbb{R}^+$. Assume that for some $\alpha_0 > 0$ and all $\alpha \in [0, \alpha_0[$,*

$$\lim_{n\to\infty} \int_0^{+\infty} \mathrm{e}^{-\alpha x}\, \mathrm{d}\pi_n(x) = \int_0^{+\infty} \mathrm{e}^{-\alpha x}\, \mathrm{d}\pi(x).$$

Then (π_n) converges weakly to π, and the convergence holds for all $\alpha \geqslant 0$.

To apply this lemma, we have to check that $(L_t(\alpha))^{1/t}$ and $\mathrm{e}^{-\ell(\alpha)}$ are the Laplace transforms of probability distributions on $\mathbb{R}^+$. It turns out that in our case, the function $L_t(\alpha)$ defined by (2) is the Laplace transform of an infinitely divisible distribution, and thus so are $(L_t(\alpha))^{1/t}$ and its limit. We give here the probabilistic interpretation of $\mathrm{e}^{-\ell_0(\alpha)}$ and $\mathrm{e}^{-\ell_1(\alpha)}$ as the Laplace transforms of two infinitely divisible distributions. Next, the particular case of a Gauss–Markov process will be considered.

Through an orthogonal transformation diagonalizing its covariance matrix, the squared norm of any Gaussian vector can be written as the sum of independent random variables, each being the square of a Gaussian variable and thus having noncentral chi-squared distribution. If Z is Gaussian with mean μ and variance v, then the Laplace transform of Z^2 is

$$\phi(\alpha) = (1 + 2\alpha v)^{-1/2} \exp\bigl(-\mu^2\alpha/(1 + 2\alpha v)\bigr).$$

The first factor is the Laplace transform of the gamma distribution with shape parameter $1/2$ and scale parameter $2v$. Assuming μ and v nonnull, rewrite the second factor as

$$\exp\left(-\frac{\mu^2}{2v}\bigl(1 - (1 + 2\alpha v)^{-1}\bigr)\right).$$

This is the Laplace transform of a Poisson compound of the exponential with expectation $2v$ by the Poisson distribution with rate $\frac{\mu^2}{2v}$. Therefore, the squared norm of a Gaussian vector has an infinitely divisible distribution, which is a convolution of gamma distributions with Poisson compounds of exponentials. Squared Gaussian vectors have received a lot of attention since even in dimension 2, the mean and covariance matrix must satisfy certain conditions for the distribution of the vector to be infinitely divisible [17]. Yet the sum of coordinates of such a vector always has an infinitely divisible distribution.

For all t, the distribution with Laplace transform $(L_t(\alpha))^{1/t}$ is the convolution of gamma distributions with Poisson compounds of exponentials. As t tends to infinity, $(L_t(\alpha))^{1/t}$ tends to $\mathrm{e}^{-\ell_0(\alpha)}\mathrm{e}^{-\ell_1(\alpha)}$. The first factor $\mathrm{e}^{-\ell_0(\alpha)}$ is the Laplace transform of a limit of convolutions of gamma distributions, which belongs to the Thorin class $T(\mathbb{R}^+)$ (see [3] as a general reference). Consider now $\mathrm{e}^{-\ell_1(\alpha)}$. Rewrite $\ell_1(\alpha)$ as

$$\begin{aligned}\ell_1(\alpha) &= m_\infty^2 \alpha\bigl(1 + 2\alpha f(0)\bigr)^{-1} \\ &= \frac{m_\infty^2}{2f(0)}\bigl(1 - \bigl(1 + 2\alpha f(0)\bigr)^{-1}\bigr).\end{aligned}$$

Thus, $\mathrm{e}^{-\ell_1(\alpha)}$ is the Laplace transform of a Poisson compound of the exponential distribution with expectation $2f(0)$ by the Poisson distribution with parameter $\frac{m_\infty^2}{2f(0)}$.

As an illustrating example, consider the Gauss–Markov process defined as follows. Let θ be a real such that $-1 < \theta < 1$. Let $(\varepsilon_t)_{t\geqslant 1}$ be a sequence of i.i.d. standard Gaussian random variables. Let Y_0, independent from the sequence $(\varepsilon_t)_{t\geqslant 1}$, follow the normal $\mathcal{N}(0, (1-\theta^2)^{-1})$ distribution. For all $t \geqslant 1$, let

$$Y_t = \theta Y_{t-1} + \varepsilon_t .$$

Thus, $(Y_t)_{t\in\mathbb{N}}$ is a stationary centered autoregressive process. Consider the noncentered process $(X_t)_{t\in\mathbb{N}}$ with $X_t = Y_t + m_\infty$. This is the case considered in [13], where a stronger result was proved. Formula (10) on p. 72 of that reference matches (4) and (5) here. Indeed, the spectral density is

$$f(\lambda) = \frac{1}{1+\theta^2 - 2\theta\cos(\lambda)} .$$

Write $\ell_0(\alpha)$ as a contour integral over the unit circle:

$$\begin{aligned}\ell_0(\alpha) &= \frac{1}{4\pi}\int_0^{2\pi} \log\bigl(1 + 2\alpha f(\lambda)\bigr)\,\mathrm{d}\lambda \\ &= \frac{1}{4\pi\mathrm{i}}\oint_{|\zeta|=1} \frac{1}{\zeta}\log\left(1 + \frac{2\alpha}{1+\theta^2-\theta(\frac{1}{\zeta}+\zeta)}\right)\mathrm{d}\zeta .\end{aligned}$$

Now we have

$$1 + \frac{2\alpha}{1+\theta^2-\theta(\frac{1}{\zeta}+\zeta)} = \frac{\zeta^2 - (\theta + \frac{1}{\theta} + \frac{2\alpha}{\theta})\zeta + 1}{\zeta^2 - (\theta+\frac{1}{\theta})\zeta + 1} .$$

Observe that the two roots of the numerator have the same sign as θ, and their product is 1. Denote them by ζ^- and ζ^+, so that $0 < |\zeta^-| < 1 < |\zeta^+|$. The two roots of the denominator are θ and $\frac{1}{\theta}$. The function to be integrated has five poles, among which $0, \theta, \zeta^-$ are inside the unit disk, and $\frac{1}{\theta}, \zeta^+$ are outside. Rewrite ℓ_0 as

$$\ell_0(\alpha) = \frac{1}{4\pi\mathrm{i}}\oint_{|\zeta|=1}\frac{1}{\zeta}\log\left(\frac{\zeta-\zeta^-}{\zeta-\theta}\right)\mathrm{d}\zeta + \frac{1}{4\pi\mathrm{i}}\oint_{|\zeta|=1}\frac{1}{\zeta}\log\left(\frac{\zeta-\zeta^+}{\zeta-\frac{1}{\theta}}\right)\mathrm{d}\zeta .$$

The first integral is null since

$$\oint_{|\zeta|=1}\frac{1}{\zeta}\log\bigl(\zeta-\zeta^-\bigr)\,\mathrm{d}\zeta = \oint_{|\zeta|=1}\frac{1}{\zeta}\log(\zeta-\theta)\,\mathrm{d}\zeta ,$$

the two functions having the same residues inside the unit disk. The second integral is

$$\frac{1}{4\pi\mathrm{i}}\oint_{|\zeta|=1}\frac{1}{\zeta}\log\left(\frac{\zeta-\zeta^+}{\zeta-\frac{1}{\theta}}\right)\mathrm{d}\zeta = \frac{1}{2}\log\bigl(\theta\zeta^+\bigr).$$

Therefore,

$$\ell_0(\alpha) = \frac{1}{2}\log\bigl(\theta\zeta^+\bigr)$$

$$= \frac{1}{2}\log\biggl(\frac{1}{2}\Bigl(\theta^2+1+2\alpha+\sqrt{\bigl((\theta+1)^2+2\alpha\bigr)\bigl((\theta-1)^2+2\alpha\bigr)}\Bigr)\biggr).$$

The expression of ℓ_1 is

$$\ell_1(\alpha) = \frac{m_\infty^2\alpha(1-\theta)^2}{(1-\theta)^2+2\alpha}.$$

It turns out that the probability distribution with Laplace transform

$$\mathrm{e}^{-\ell_0(\alpha)} = \biggl(\frac{1}{2}\Bigl(\theta^2+1+2\alpha+\sqrt{\bigl((\theta+1)^2+2\alpha\bigr)\bigl((\theta-1)^2+2\alpha\bigr)}\Bigr)\biggr)^{-1/2}$$

has an explicit density $f_0(x)$ defined on $(0,+\infty)$, which is related to the modified Bessel function of the first kind with order $1/2$ (compare with formula (3.10) on p. 437 in [7]):

$$\begin{aligned} f_0(x) &= \mathrm{e}^{-\frac{1+\theta^2}{2}x}\bigl(2^{-1}|\theta|^{-1/2}x^{-1}I_{1/2}(|\theta|x)\bigr) \\ &= \mathrm{e}^{-\frac{1+\theta^2}{2}x}\bigl((2\pi)^{-1/2}|\theta|^{-1}x^{-3/2}\sinh(|\theta|x)\bigr). \end{aligned}$$

References

[1] Benitz, G.R., Bucklew, J.A.: Large deviation rate calculations for nonlinear detectors in Gaussian noise. IEEE Trans. Inf. Theory **36**(2), 358–371 (1990). MR1052787. doi:10.1109/18.52482

[2] Bingham, N.H.: Szegő's theorem and its probabilistic descendants. Probab. Surv. **9**, 287–324 (2012). MR2956573

[3] Bondesson, L.: Generalized Gamma Convolutions and Related Classes of Distributions. Lect. Notes Stat., vol. 76. Springer (1992). MR1224674. doi:10.1007/978-1-4612-2948-3

[4] Böttcher, A., Silbermann, B.: Introduction to Large Truncated Toeplitz Matrices. Springer (1999). MR1724795. doi:10.1007/978-1-4612-1426-7

[5] Bryc, W., Dembo, A.: Large deviations for quadratic functionals of Gaussian processes. J. Theor. Probab. **10**, 307–332 (1997). MR1455147. doi:10.1023/A:1022656331883

[6] Dembo, A., Zeitouni, O.: Large Deviations Techniques and Applications, 2nd edn. Springer (1998). MR1619036. doi:10.1007/978-1-4612-5320-4

[7] Feller, W.: An Introduction to Probability Theory and Its Applications vol. II, 2nd edn. Wiley, London (1971). MR0270403

[8] Gray, R.M.: Toeplitz and circulant matrices: A review. Found. Trends Commun. Inf. Theory **2**(3), 155–239 (2006)

[9] Grenander, U., Szegő, G.: Toeplitz Forms and Their Applications, 2nd edn. Chelsea, New York (1984). MR0890515

[10] Hannan, E.J.: Multiple Time Series. Wiley, New York (1970). MR0279952

[11] Kallenberg, O.: Foundations of Modern Probability. Springer (1997). MR1464694

[12] Kleptsyna, M.L., Le Breton, A., Viot, M.: New formulas concerning Laplace transforms of quadratic forms for general Gaussian sequences. J. Appl. Math. Stoch. Anal. **15**(4), 309–325 (2002). MR1950568. doi:10.1155/S1048953302000266

[13] Kleptsyna, M.L., Le Breton, A., Ycart, B.: Exponential transform of quadratic functional and multiplicative ergodicity of a Gauss–Markov process. Stat. Probab. Lett. **87**, 70–75 (2014). MR3168937. doi:10.1016/j.spl.2013.12.023

[14] Kontoyiannis, I., Meyn, S.P.: Spectral theory and limit theorems for geometrically ergodic Markov processes. Ann. Appl. Probab. **13**(1), 304–362 (2003). MR1952001. doi:10.1214/aoap/1042765670

[15] Krein, M.G.: Integral equations on the half-line with a kernel depending on the difference of the arguments. Usp. Mat. Nauk **13**(5), 3–120 (1958). MR0102721

[16] Louhichi, S., Ycart, B.: Exponential growth of bifurcating processes with ancestral dependence. Adv. Appl. Probab. **47**(2), 545–564 (2015). MR3360389. doi:10.1239/aap/1435236987

[17] Marcus, M.B., Rosen, J.: Existence of a critical point for the infinite divisibility of squares of Gaussian vectors in R^2 with non-zero mean. Electron. J. Probab. **14**(48), 1417–1455 (2009). MR2519526

[18] Muskhelishvili, N.I.: Singular Integral Equations: Boundary Problems of Function Theory and Their Application to Mathematical Physics, 3rd edn. Dover, New York (2008). MR1215485

[19] Yurinsky, V.: Sums and Gaussian Vectors. Lect. Notes Math., vol. 1617. Springer (1995). MR1442713